MIR503HG Loss Promotes Endothelial-to-Mesenchymal Transition in Vascular Disease

MIR503HG Loss Promotes Endothelial-to-Mesenchymal Transition in Vascular Disease

João P. Monteiro[#1], Julie Rodor[#1], Axelle Caudrillier[1], Jessica P. Scanlon[1], Ana-Mishel Spiroski[1], Tatiana Dudnakova[1], Beatrice Pflüger-Müller[2,3], Alena Shmakova[1], Alex von Kriegsheim[4], Lin Deng[1], Richard S. Taylor[5], John R. Wilson-Kanamori[6], Shiau-Haln Chen[1], Kevin Stewart[1], Adrian Thomson[1], Tijana Mitić[1], John D. McClure[7], Jean Iynikkel[1], Patrick W.F. Hadoke[1], Laura Denby[1], Angela C. Bradshaw[7], Paola Caruso[8], Nicholas W. Morrell[8], Jason C. Kovacic[9,10], Igor Ulitsky[11], Neil C. Henderson[6], Andrea Caporali[1], Matthias S. Leisegang[2,3], Ralf P. Brandes[2,3], Andrew H. Baker[1*]

[1]The Queen's Medical Research Institute, Centre for Cardiovascular Science, University of Edinburgh; [2]Institute for Cardiovascular Physiology, Goethe University; [3]German Center of Cardiovascular Research (DZHK), Partner site RheinMain, Frankfurt, Germany. [4]Edinburgh Cancer Research UK Centre, Institute of Genetics and Molecular Medicine, University of Edinburgh; [5]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh; [6]The Queen's Medical Research Institute, Centre for Inflammation Research, University of Edinburgh; [7]Institute of Cardiovascular and Medical Sciences, BHF Glasgow Centre, University of Glasgow; [8]BHF Cambridge (CRE), University of Cambridge; [9]The Zena and Michael A. Wiener Cardiovascular Institute, School of Medicine at Mount Sinai, New York; [10]Victor Chang Cardiac Research Institute, Darlinghurst, Australia; St Vincent's Clinical School, University of NSW, Australia, and;[11]Biological Regulation, Weizmann Institute of Science, Rehovot, Israel.

[#]Equal contribution, *Corresponding author

Running Title: Loss of MIR503HG Promotes EndMT During Disease

Subject Terms:
Basic Science Research
Cell Signaling/Signal Transduction
Cellular Reprogramming
Endothelium/Vascular Type/Nitric Oxide
Vascular Biology

Address correspondence to:
Dr. Andrew H. Baker
University of Edinburgh
Queen's Medical Research Institute
47 Little France Crescent
Edinburgh
UK, EH16 4TJ
Tel: +44 131 24 26728
Andy.Baker@ed.ac.uk

ABSTRACT

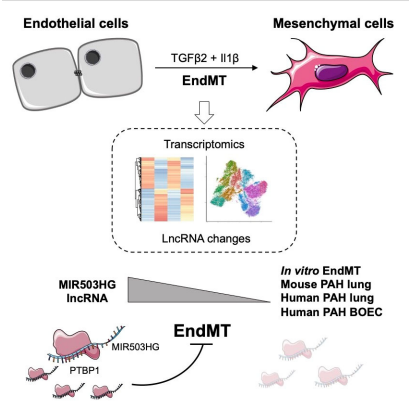
Rationale: Endothelial-to-mesenchymal transition (EndMT) is a dynamic biological process involved in pathological vascular remodelling. However, the molecular mechanisms that govern this transition remain largely unknown, including the contribution of long non-coding RNAs (lncRNAs).

Objectives: To investigate the role of lncRNAs in EndMT and their relevance to vascular remodelling.

Methods and Results: To study EndMT *in vitro*, primary endothelial cells (EC) were treated with transforming growth factor-β2 and interleukin-1β. Single-cell and bulk RNA-sequencing were performed to investigate the transcriptional architecture of EndMT and identify regulated lncRNAs. The functional contribution of seven lncRNAs during EndMT was investigated based on a DsiRNA screening assay. The loss of lncRNA *MIR503HG* was identified as a common signature across multiple human EC types undergoing EndMT *in vitro*. *MIR503HG* depletion induced a spontaneous EndMT phenotype, while its overexpression repressed hallmark EndMT changes, regulating 29% of its transcriptome signature. Importantly, the phenotypic changes induced by *MIR503HG* were independent of *miR-424* and *miR-503*, which overlap the lncRNA locus. The pathological relevance of *MIR503HG* down-regulation was confirmed *in vivo* using Sugen/Hypoxia (SuHx)-induced pulmonary hypertension (PH) in mouse, as well as in human clinical samples, in lung sections and blood outgrowth endothelial cells (BOECs) from pulmonary arterial hypertension (PAH) patients. Overexpression of human *MIR503HG* in SuHx mice led to reduced mesenchymal marker expression, suggesting MIR503HG therapeutic potential. We also revealed that *MIR503HG* interacts with the Polypyrimidine Tract Binding Protein 1 (PTB1) and regulates its protein level. PTBP1 regulation of EndMT markers suggests that the role of *MIR503HG* in EndMT might be mediated in part by PTBP1.

***Conclusions*:** This study reports a novel lncRNA transcriptional profile associated with EndMT and reveals the crucial role of the loss of *MIR503HG* in EndMT and its relevance to pulmonary hypertension.

Keywords:
Endothelium, endothelial-to-mesenchymal transition, EndMT, lncRNA, miRNA, endothelial cell, endothelial cell differentiation, vascular remodeling, molecular biology, non-coding RNA.



Nonstandard Abbreviations and Acronyms:

BOEC: blood outgrowth endothelial cells
EC: Endothelial cell
EndMT: Endothelial to mesenchymal transition
GO: Gene ontology
HUVEC: Human umbilical vein endothelial cells
HPAEC: Human pulmonary artery endothelial cells
IL-1β: interleukin-1β
LncRNA: long non-coding RNA
MiRNA/Mi-R: MicroRNA
PAH: Pulmonary arterial hypertension
PH: Pulmonary hypertension
(sc)RNA-seq: (Single-Cell) RNA sequencing
SuHx: Sugen Hypoxia PH mouse model
TGF-β: Transforming growth factor-β2

INTRODUCTION

Endothelial to mesenchymal transition (EndMT) is a complex cellular transdifferentiation process whereby endothelial cells (EC) lose their endothelial identity and acquire mesenchymal cell characteristics[1]. Akin to epithelial-to-mesenchymal transition, EndMT has been described during heart embryonic development[1], as well as in neovascularization and tissue repair[2]. While its role remains under debate, EndMT has been shown to be involved in a variety of maladaptive tissue remodelling scenarios[3-6] such as atherosclerosis[7], and vein graft remodelling[8]. Such remodelling is also causal in pulmonary arterial hypertension (PAH), a progressive disease with limited treatment options, largely defined by structural alterations to the lung vasculature. These alterations include stiffening of proximal pulmonary arteries, increased intimal and medial arterial thickness, along with the eventual development of complex neointimal lesions. Underlying many of these changes is the increased deposition of extracellular matrix and appearance of smooth muscle-like cells with high proliferative and migratory potential[9]. Several reports now suggest the involvement of EndMT during both the development and progression of PAH, opening new therapeutic avenues[10-12]. As both PAH and EndMT are linked with imbalance in TGF-β signaling, several *in vitro* models have been developed relying on TGF-β signaling as the main inducer of EndMT, alone or with an additional stimulus[7, 13].

EndMT is characterised by the decrease of endothelial markers such as PECAM1/CD31 and VE-Cadherin (CDH5) in association with the gain of mesenchymal markers such as α-smooth muscle actin (ACTA2), calponin (CNN1)[1], as well as enhanced collagen production. EndMT is also associated with increased expression of several transcription factors such as TWIST, SMAD3, SNAI1 and SNAI2[1]. However, the precise molecular mechanisms and upstream determinants governing this transition, remain to be rigorously defined, particularly when considering vascular heterogeneity and the presence of distinct stages of EndMT. Indeed, cells that have lost EC markers and thus undergone a complete EndMT process showed high proliferative and migratory capacity whereas partial EndMT cells, still expressing EC markers, have also been isolated[12]. While both protein-coding genes and miRNAs have been widely involved in EndMT regulation[1, 14], the role of long non-coding RNAs (lncRNAs) is yet to be thoroughly explored. LncRNAs are long (>200nt) non-coding transcripts, often expressed at low levels but with high specificity. Several lncRNAs have been described as critical regulators of gene expression both at the transcriptional and post-transcriptional level and have roles in a wide range of diseases[15]. LncRNAs have already been associated with EC function and dysfunction[16]. The lncRNA *MALAT1* (metastasis associated lung

adenocarcinoma transcript 1) has been implicated in the modulation of TGF-β1-induced EndMT through regulation of the *miR-145*-TGFBR2/SMAD3 axis[17]. Recently, the lncRNA *GATA6-AS* (antisense transcript of *GATA6*) was shown to suppress TGF-β2-induced EndMT *in vitro* via targeting LOXL2 (Lysyl oxidase homolog 2)[18].

Here, we have applied an unbiased approach using single-cell and bulk RNA sequencing of an *in vitro* model of EndMT to characterise EndMT transcriptional changes. We identified 103 differentially expressed lncRNAs during EndMT progression in both venous and arterial ECs. Among these lncRNAs, the substantive loss of the lncRNA *MIR503HG* was a consistent molecular event during EndMT, both *in vitro* and *in vivo* mouse and human samples. We further showed a direct contribution of *MIR503HG* on the regulation of EndMT, reporting large transcriptional changes and preliminary data on its therapeutic potential.

METHODS

The authors declare that all methods, material list and processed experimental data are available within the article and its online files. Any additional data are available from the corresponding author upon request. Raw and processed sequencing data are available at GEO database (accession: GSE118446, GSE118815, GSE159843). All methods are included in the Online Supplemental Materials. Please see the Major Resources Table for material list.

RESULTS

Induction of EndMT in arterial and venous ECs using TGF-β2 and IL-1β.

To replicate EndMT *in vitro* and characterise its molecular signature, human primary venous (HUVEC) and arterial (HPAEC) endothelial cells were subjected to a combination of transforming growth factor beta 2 (TGF-β2) and interleukin 1 beta (IL-1β) for 7 days[13]. RT-qPCR analysis confirmed EndMT induction in TGF-β2 and IL-1β co-treated cells, but not TGF-β2 or IL-1β alone, with the loss of the endothelial marker *PECAM1*, gain of mesenchymal markers *ACTA2* and *COL1A1* and the EndMT-associated transcription factor *SNAI2* (Figure 1A and Online Figure I_A). In addition, changes to PECAM1 and SNAI2 expression were confirmed at the protein level by immunofluorescence microscopy (Figure 1B and Online Figure I_B). We also confirmed the decrease of PECAM1 and increase of ACTA2 protein levels by Western blot in HPAEC (Online Figure I_C). We also assessed the effect of the TGF-β2 and IL-1β co-treatment on the proliferative and migratory capacities of HUVEC and observed a significant decrease of cellular proliferation (Figure 1D) but no significant change in the migratory capacity of EndMT cells (Figure 1E).

Together, these results confirm the induction of an EndMT-like profile in both arterial and venous EC using a TGF-β2 and IL-1β co-treatment model.

Transcriptional profiling of primary ECs undergoing EndMT.

To characterise EndMT transitioning populations and transcriptomic changes associated with this transition, we carried out single-cell RNA sequencing (scRNA-seq) on co-treated and control HUVEC at day 0, day 3 and day 7 using 10X Genomics technology. Based on tSNE dimensionality reduction, we confirmed the clear separation between control and TGF-β2 and IL-1β co-treated cells (Figure 2A). While

day 3 and day 7 untreated cells showed some overlap, day 7 treated cells clustered separately from day 3 treated cells showing a time-dependent effect of the treatment (Figure 2A). The loss of an endothelial signature was confirmed using *PECAM1*, *CDH5*, *ICAM2*, *ERG* and *VWF*, and the gain of a mesenchymal/EndMT signature using *TAGLN*, *COL1A1*, and *S100A4* and *SNAI2* expression (Figure 2B). Eleven clusters were identified (Figure 2C), with clusters 1-6 corresponding to control HUVEC and displaying an endothelial transcriptional program, whereas clusters 7-11 from co-treated HUVEC displayed a downregulation of endothelial genes and progression to a mesenchymal signature (Figure 2D). In agreement with the decrease of proliferation described above (Figure 1D), we observed a decrease of *MKI67* expression in groups 8-11 (Figure 2D). Cluster 11, composed exclusively from day 7 co-treated HUVEC, constituted a new 'advanced EndMT' population (Figure 2D). Cluster 11 was characterised by the expression of 132 marker genes, including known EndMT markers (*TGF□R1*, *NOTCH2* and *BMP2*) and novel genes involved in cell differentiation (*ANKRD1*, *OSGIN2*) and stress signaling (*CXCL8* and *CSF3*) (Online Figure II_A-B).

To identify the complete transcriptomic changes associated with EndMT, including lncRNAs, we performed high-depth bulk RNA-sequencing of HUVEC and HPAEC undergoing EndMT after 7 days (Figure 2, Online Figure III_A). Principal component analysis (PCA) confirmed a clear segregation of the different treatment groups, compared to untreated controls (Figure 2E, Online Figure III_B). HUVEC and HPAEC derived EndMT-cells clustered separately from their respective control group, but also from each other, suggesting transcriptional heterogeneity in the induction of EndMT across vascular beds (Figure 2E, Online Figure III_B). We identified 1721 up-regulated and 1260 down-regulated genes in EndMT-HUVEC compared to untreated cells (Online Dataset I), with 46% of the up-regulated and 60% of the down-regulated genes not observed in the single-treatment groups (Figure 2F, Online Figure III_C-D). Importantly, around 40% of regulated genes in HUVECs were also affected in HPAEC (Figure 2F-G and Online Dataset I). Gene Ontology (GO) analysis found that up-regulated genes associated with immune response, while down-regulated genes associated with GO terms related to DNA conformation and cell cycle (Online Figure III_E and Online Dataset II).

LncRNA expression was affected upon EndMT induction, with 69 lncRNAs up-regulated and 34 down-regulated in HUVEC and HPAEC (Online Dataset I). Top lncRNA candidates were selected based on fold change, level of expression and genomic location: *MIR503HG*, *HOTAIRM1*, *AC123023.1*, *CTC-378H22.1*, *LINC00702*, *MIR3142HG*, *RP11-37B2.1*, *AC147651.4* and *RP11-79H23.3* (Figure 2H and Online Figure IV_A-B). Differential expression patterns were confirmed by RT-qPCR in HUVEC and HPAEC (Online Figure IV_C-D) for all candidate lncRNAs except *CTC-378H22.1* and *HOTAIRM1*, which failed validation and were removed from downstream analysis (Online Figure IV_A-B). Using the single-cell RNA-seq dataset, we confirmed the expression changes of all candidates in the 'advanced' EndMT population (Figure 2I). Interestingly, the change for some candidates, including *MIR503HG*, was also observed in the day 3 sample clusters (Figure 2I) suggesting a role in the early phase of the transition.

Screening for lncRNA function during EndMT.

A siRNA-mediated knockdown approach was utilised to screen the functional contribution of each lncRNA to EndMT (Online Figure V) based on the expression of defined EndMT markers (i.e. *PECAM1*, *ACTA2*, *SNAI2*, *COL1A1*). Significant knockdowns were confirmed for four lncRNAs but not for LINC00702, *AC147651.4* and *AC123023.1*, which were subsequently eliminated from this study (Online Figure V_A). Targeted depletion of the selected lncRNA candidates lead to significant changes in EndMT markers, highlighting their potential causal contribution to EndMT (Online Figure V_B). Notably, knockdown of *MIR503HG* induced a robust EndMT profile, even in the absence of TGF-β2 and IL-1β co-treatment (Online Figure V_A-B), and as such was chosen for further analysis.

Loss of MIR503HG is a common feature of EndMT in vitro.

MIR503HG is an intergenic lncRNA located on chromosome X with five reported isoforms on GENCODE.v26 (Figure 3A and Online Figure VI_A). While RNA-seq data showed a significant down-regulation of all isoforms (Online Figure VI_B), we confirmed the down-regulation of the top three expressed isoforms 2, 3 and 5 by RT-qPCR following EndMT induction in HUVEC (Online Figure VI_B-C). Importantly, as the *MIR503HG* gene locus overlaps with the *miR-424* and *miR-503* genes (Figure 3A), we analysed their expression during EndMT and observed a significant down-regulation of both *miR-424-5p* and *miR-503-5p* (Figure 3B).

Interestingly, the final 595 base-pair region of isoform 2 (*MIR503HG_2*) is highly conserved based on PhyloP score (Online Figure VI_A), as was its secondary substructure[19]. As such, we chose *MIR503HG_2* as our main transcript of interest. In addition to HUVEC and HPAEC (online Figure IV_C-D), down-regulation of *MIR503HG_2* after TGF-β2 and IL-1β co-stimulation was also seen in Human Saphenous Vein Endothelial Cells (HSVEC) and Human Coronary Artery Endothelial Cells (HCAEC) (Figure 3C). Additionally, *MIR503HG_2* expression was decreased in a second *in vitro* model of EndMT using TGF-β2 and H_2O_2 co-stimulation[7] (Figure 3D). Using our single-cell RNA-seq data, we also confirmed the positive correlation of *MIR503HG* with the EC signature and a negative correlation for mesenchymal signature (Figure 3E).

As subcellular localisation of lncRNA can provide information regarding function, *MIR503HG* localisation was determined by RT-qPCR of nuclear/cytoplasmic fraction (Figure 3F) and RNA-FISH (Figure 3G) in HUVEC. Together, this data revealed that *MIR503HG* is predominantly localised to the nucleus, with a mean of 2.16 copies per nuclei (Figure 3G).

Loss of MIR503HG initiates EndMT in the absence of TGF-β2 and IL-1β.

The siRNA screening approach was validated on a larger sample size. We confirmed that all expressed *MIR503HG* isoforms were down-regulated in the si503HG samples (Online Figure VII_A) and that, in the absence of treatment, si503HG induces an EndMT profile based on RT-qPCR analysis (Figure 4A-B). These effects were mirrored at the protein level based on immunofluorescence (Figure 4C) and western blot (Online VII_B). Furthermore, we showed that *MIR503HG* depletion replicated the key phenotypes of the EndMT model, such as loss of EC monolayer integrity (Figure 4C), decreased proliferation (Online Figure VII_C) but also showed a decrease of cell migration (Online Figure VII_D). Given the nuclear localisation of *MIR503HG*, an antisense GapmeR (gap503HG) was also used to manipulate *MIR503HG* expression (Online Figure VIII_A) and replicated the findings of the siRNA-mediated knockdown, inducing similar expression changes to EndMT markers at the RNA (Online Figure VIII_B) and protein levels (Online Figure VIII_C).

We observed a significant reduction in the expression of *miR-424* and *miR-503* with siRNA but not GapmeR-mediated MIR503HG depletion (Online Figure IX), presumably due to the oligonucleotides targeting the miRNA precursor. To target specifically *MIR503HG*, we used a lentiviral CRISPR/Cas9 gene editing system to delete the conserved exon of *MIR503HG* locus without disturbing the miRNA cluster upstream. Co-transduction of HUVEC using two separate lentiviral CRISPR/Cas9 guide RNA (gRNA) pairs, led to a >50% reduction in the expression of the target region (exon 3) (Figure 4D) while maintaining *miR-424* and *miR-503* as well as *MIR503HG* exon 1 expression after 8 days (Figure 4D-E). As with our previous knockdown strategies, reduced *MIR503HG* availability in HUVEC resulted in similar changes to EndMT marker expression (Figure 4F). Of note, for one of the guide combinations, no statistical difference was observed for *PECAM1* and *SNAI2* expression. Finally, an anti-miR strategy was used to interrogate the direct contribution of *miR-503* and *miR-424* to EndMT in HUVEC. Knocking down either *miR-424* or *miR-503* (Online Fig X_A) had no significant effect on the expression level of *MIR503HG* (Online Figure X_B).

Moreover, EndMT was not induced by knockdown of either miRNA, with no statistical difference in endothelial or mesenchymal markers observed (Online Figure X_C). These data showed that the role of *MIR503HG* on EndMT regulation is independent of *miR-424* and *miR-503*.

Overexpression of MIR503HG isoform 2 represses EndMT.

To assess if *MIR503HG* overexpression could prevent EndMT, we designed a lentiviral vector carrying the *MIR503HG_2* transcript sequence (LNT_503HG). HUVEC were transfected with LNT_503HG or LNT_CT (lentivirus control). At 3 days post transduction, LNT_503HG samples showed an increase of *MIR503HG_2* only and not isoforms 3 and 5, when compared to LNT_CT (Online Figure XI_A). Cellular fractionation and RNA-FISH confirmed the predominant nuclear localization of overexpressed *MIR503HG,* with a mean number of 6.46 copies per nuclei (Online Figure XI_B-D). Overexpression of *MIR503HG* (Figure 5A) did not affect expression of endothelial or mesenchymal markers under control conditions (Figure 5B). However, under EndMT conditions, *MIR503HG* overexpressing cells (Figure 5A) showed an increase in *PECAM1* expression, and a marked suppression of *SNAI2* and *COL1A1* expression compared to LNT_CT (Figure 5B). Immunofluorescence validated these results at the protein level (Figure 5C). Overexpression of *MIR503HG* did not affect *miR-424* or *miR-503* levels compared to LNT_CT, in either control or EndMT-conditions (Online Figure XII), further confirming a miRNA-independent role of MIR503HG on EndMT.

To identify the contribution of *MIR503HG* overexpression on the EndMT transcriptomic changes, we performed RNA-seq on control and EndMT HUVEC with lentiviral overexpression (Online Dataset III). PCA revealed that LNT_503HG_2 had no major effect on the transcriptome of untreated cells (Control_LNT_503HG_2), clustering with Control and Control_LNT_CT cells (Online Figure XIII). Interestingly, EndMT_LNT_503HG cells clustered in between Control and EndMT cells (Online Figure XIII), suggesting the treatment with LNT_503HG is preventing the EndMT process. We identified 803 and 880 genes up- and down-regulated, respectively, by *MIR503HG* overexpression in EndMT (EndMT_LNT_503HG compared to EndMT_LNT_CT) (Figure 5D). We found a 29% overlap between EndMT regulated genes and the genes regulated by LNT_503 in treated conditions (Figure 5E). These overlapping genes include additional markers of EndMT such as the endothelial genes *NR2F2* and *NOS3* as well as the mesenchymal genes *TAGLN*, *FN1* and *CNN1*. This data again confirms the strong contribution of *MIR503HG* to EndMT.

MIR503HG expression is lost during vascular remodelling in a mouse PH model.

To examine the contribution of *MIR503HG in vivo*, we used a Sugen/Hypoxia (SuHx) PH mouse model, in which vessel remodelling involves EndMT[10] (Figure 6A). Additionally, we used an inducible endothelial lineage tracing system of Ind.Endotrack (Cdh5-CreER^{T2}-TdTomato) mice[20, 21], in which TdTomato is constitutively expressed after Tamoxifen treatment in all EC regardless of subsequent changes in cellular phenotype[8]. Ind.Endotrack mice were subjected to SuHx or control conditions for 3 weeks. TdTomato^{+} cells isolated from SuHx lungs presented an EndMT profile with increased expression of mesenchymal-specific markers (*Acta2*, *Vimentin* and *Col1a1, Snai2*) and loss of endothelial specificity (*Pecam1* and *Cdh5*) when compared to control mice (Figure 6B). Notably, expression of the *MIR503HG* mouse homolog, *Gm28730* (locus and primers shown in Online Figure XIV), was found to be significantly reduced in SuHx TdTomato^{+} EC (Figure 6C).

To establish the role of *MIR503HG* in EndMT *in vivo*, we overexpressed the human *MIR503HG_2* transcript in mice and assessed its impact on the EndMT markers after PH induction. We optimised the uptake of a GFP lentiviral construct on normoxic mice, using intranasal delivery. Flow cytometry analysis confirmed the presence of GFP^{+} cells in the lung, 13.7% of which corresponded to CD31^{+} endothelial cells (Online Figure XV_A). This delivery protocol was utilised to deliver either LNT_503HG or LNT_CT to 8

to 10-week-old C57BL/6 mice, which then underwent PH induction by SuHx (Figure 6D). Endothelial ($CD31^+$) and $CD31^-$ lung cells were isolated by flow cytometry (Online Figure XV_B) and the expression of *MIR503HG* and EndMT markers was assessed. Human MIR503HG_2 expression was up-regulated in both $CD31^-$ and $CD31^+$ lung cells for LNT_503HG mice compared to LNT_CT (Figure 6E and Online Figure XVI_A). The mouse MIR503HG homolog was also increased in LNT_503HG versus LNT_CT $CD31^+$ ECs (Figure 6E). Importantly, analysis of EndMT markers revealed the significant down-regulation of the mesenchymal markers *Acta2* and *Col1a1* in LNT_503HG versus LNT_CT $CD31^+$ ECs but not in $CD31^-$ cells (Figure 6F and Online Figure XVI_B).

These experiments show that the down-regulation of *MIR503HG* mouse homolog *Gm28730* in PH is associated with EndMT and that human *MIR503HG_2* overexpression can prevent the induction of mesenchymal markers in the lung endothelial compartment in response to EndMT-inducing stimuli *in vivo*.

Loss of MIR503HG is associated with EndMT in Human PAH.

Blood outgrowth ECs (BOECs) derived from PAH patients have been shown to recapitulate pulmonary endothelial cell dysfunction[22]. Based on the expression of endothelial and mesenchymal markers, we showed that BOECs derived from PAH patients presented an EndMT-like phenotype compared to cells from control patients that was accompanied by a significant reduction of *MIR503HG_3* and *MIR503HG_5* expression (Figure 7A).

To determine the association between the loss of *MIR503HG* and vascular remodelling, we analysed lung tissue sections from three PAH patients and three control lungs, using *in situ* hybridisation. In control lungs, *MIR503HG* expression was present throughout the vasculature (Figure 7B-C, Online Figure XVII_A-B and Online Figure XVIII_A). Conversely, in PAH lung vasculature, *MIR503HG* expression was detected in non-remodelled vessels but was absent from remodelled arterial vessels (Figure 7B-C, Online Figure XVII_C-D and VXIII).

Together, this demonstrated that loss of *MIR503HG* is observed in PAH patients in association with vascular remodelling.

MIR503HG regulation of EndMT is mediated in part by PTBP1.

To study the *MIR503HG* mechanism of action, we performed an RNA pulldown assay using *in vitro* synthesized biotinylated *MIR503HG_2* transcripts, which were captured on streptavidin beads and incubated with HUVEC extracts depleted from the cytosolic fraction (Figure 8A and online Figure XIX_A-B). Proteomic analysis by mass spectrometry (MS) was used to identify 125 significantly enriched proteins bound to biotinylated *MIR503HG_2* compared to an eGFP RNA control (Online Dataset IV). The top two enriched proteins (ranked based on fold change over control) were the RNA splicing regulatory protein Polypyrimidine tract-binding protein 1 (PTBP1) and the heterogeneous nuclear ribonucleoprotein A0 (HNRNPA0) (Figure 8B). Interaction of *MIR503HG_2* with either PTBP1 or HNRNPA0 was validated by reciprocal RNA immunoprecipitation (RIP). *MIR503HG_2* RNA showed a high and significant enrichment for PTBP1 and HNRNPA0 pulldown compared to an IgG control pulldown, while, in comparison, *UBC* mRNA showed a lower enrichment (Figure 8C).

To understand the consequences of *MIR503HG* interaction with PTBP1 and HNRNPA0 proteins, we analysed PTBP1 and HNRNPA0 protein levels in *MIR503HG* knockdown samples. Upon *MIR503HG* depletion, we observed a significant down-regulation of PTBP1 (Figure 8D) and HNRNPA0 (online Figure XIX_C). Similarly, PTBP1 protein down-regulation also occurs in EndMT (Figure 8E), along with HNRNPA0 (online Figure XIX_D). To determine if PTBP1 or HNRNPA0 down-regulation contribute to EndMT, we performed siRNA-mediated *PTBP1* and *HNRNPA0* knockdown in HUVEC (Online Figure

XIX_E-F) and assessed EndMT markers after 7 days. Interestingly, *PTBP1* depletion led to a significant decrease of *PECAM1* as well as the up-regulation of *COL1A1* and *ACTA2*, although *SNAI2* was unaffected (Figure 8F), without significantly affecting *MIR503HG* levels (Online Figure XIX_G-H). In contrast, *HNRNPA0* depletion only led to a significant increase of *ACTA2* (Figure 8F). Furthermore, we mined publicly available RNA-seq datasets from *PTBP1* knockdown in HepG2 cells, a liver hepatocellular carcinoma with epithelial cell morphology, and compared the differentially expressed genes with LNT503 regulated genes. This revealed a significant overlap between genes negatively regulated by *MIR503HG* and PTBP1 (online Figure XX), with 79 genes also up-regulated during EndMT (Figure 8G). Of interest, the overlapping genes include several known mesenchymal markers: *TAGLN*, *COL1A1*, *CNN1*, *VCAN* and *MYL9*. Altogether, these data suggest that the induction of EndMT by *MIR503HG* depletion may be mediated in part by a reduction in PTBP1 level and consequent upregulation of mesenchymal markers.

DISCUSSION

With the involvement of EndMT in pathological vascular remodelling[11, 23, 24], the in-depth characterisation of this process and its therapeutical targeting seems crucial. Here, we report the transcriptomic profile of primary EC undergoing EndMT *in vitro* using single-cell and bulk RNA-seq. We provide evidence of lncRNA expression changes associated with EndMT and the effect of the depletion of four candidate lncRNAs on EndMT. We have shown that loss of the lncRNA *MIR503HG* is pivotal to the induction of EndMT, both *in vitro* and *in vivo,* by a mechanism of action independent of *miR-424* and *miR-503*. *MIR503HG* overexpression in PAH mice affects mesenchymal marker expression, highlighting potential for therapeutic intervention. Furthermore, *MIR503HG* loss was associated with human PAH, further confirming its clinical relevance. Our data revealed the interaction of *MIR503HG* with the RNA binding protein PTBP1 and suggests PTBP1 protein regulation by *MIR503HG* contributes to EndMT. A graphical summary of this mechanism is shown in Figure 8H.

EndMT has been observed in diseases affecting different vascular beds, from PAH[23, 24] to atherosclerosis[25] and vein graft remodelling[8, 12]. While context specific regulations of EndMT likely occur, common markers and pathways have been described[1]. Our data showing a large overlap of transcriptional changes between HUVEC and HPAEC undergoing EndMT support some common mechanisms, which could be targeted as a therapeutic strategy across diseases. Our work on *MIR503HG* revealed its relevance to EndMT in several EC types, two *in vitro* EndMT models, in a mouse model of PH and in human PAH samples.

We propose that the TGF-β2 and IL-1β co-treatment model replicates an early stage of EndMT, as co-treated ECs still express the endothelial marker PECAM1 after 7 days, despite a significant down-regulation of its expression. At this time point, we showed a decrease of cell proliferation and no statistical change in migration. While fully transitioned cells have increased proliferative and migratory capacity[12], this might not be the case for transitioning cells. Indeed, decreased proliferation has been observed in another *in-vitro* EndMT model[7, 10] and similar phenotypes have been described previously in primary ECs upon activation of the TGF-β pathway[26]. Further work is required to characterise the different stages of EndMT, in terms of gene expression as well as cell phenotypic properties, both *in vitro* and *in vivo*.

MIR503HG was initially described as a hypoxia sensitive lncRNA in EC[27] and several studies have independently focused on its role in cancer. While *MIR503HG* generally inhibits cell proliferation, invasion and migration in cancer cells [19, 28-31], except in lymphoma[32], silencing of *MIR503HG* in ECs has been shown to decrease proliferation and migration[27], corroborating our findings. These observations highlight the cell type specific effect of lncRNA modulation. Our study is the first to provide evidence for the involvement of *MIR503HG* in EndMT.

Our study provides mechanistic data on *MIR503HG*, identifying binding partners in HUVEC. *MIR503HG* interaction with PTBP1 and HNRNPA0 was confirmed by reciprocal RNA immunoprecipitation. Several lncRNAs have been shown to interact with PTBP1, regulating its role in transcription[33], splicing[34] or RNA stability[35]. PTBP1 protein level was decreased in *MIR503HG* knockdown conditions and EndMT, suggesting *MIR503HG* interaction with PTBP1 in HUVEC may promote its stability. Knockdown of *PTBP1* in HUVEC and HepG2 cells led to the upregulation of mesenchymal markers, showing PTBP1 plays a role in maintaining cell identity. The interaction of *MIR503HG* with HNRNPA2B1, observed in hepatocellular carcinoma[31], was not detected in the HUVEC pulldown, suggesting cell type specific mechanism of *MIR503HG*. Collectively, we propose that the role of *MIR503HG* in EndMT is mediated in part by its regulation of PTBP1 protein level.

Although *miR-424* and *miR-503* have roles in epithelial-to-mesenchymal transition[36, 37], our data shows that *MIR503HG* modulation but not miRNA modulation, has an effect on the EndMT process. Indeed, CRISPR/Cas9 mediated deletion of the *MIR503HG* locus as well as *MIR503HG* transcript overexpression was sufficient to regulate EndMT, without affecting either miRNA expression. Crucially, targeted depletion of *miR-503* or *miR-424* did not show a statistical difference in the expression of EndMT markers. A decrease of *miR-503* and *miR-424* during EndMT or upon siRNA mediated-*MIR503HG* depletion is likely due to the overlap between the miRNA precursor and *MIR503HG* and their potential transcriptional co-regulation. Whilst many miRNA host lncRNAs have been described, for the large part their function as independent lncRNAs, instead of pri-miRNAs, have not been determined, with the exception of *MIR205HG/LEADeR*[38] and *MIR100HG*[39]. *LEADeR* lncRNA is produced through splicing of the MIR205HG locus and regulates differentiation of human prostate basal cells independently of *miR-205* function[38], while *MIR100HG* regulates cell cycle through its interaction with the RNA binding protein HuR[39]. Our data provide another example of miRNA host gene with a long non-coding RNA role, highlighting the need to dissociate the effect of host genes from their overlapping miRNAs.

Our study also shows a clear correlation between loss of MIR503HG and the gain of EndMT markers, using a mouse model of PH and human BOECs. In PAH lung tissues, we showed the association between *MIR503HG* loss and vessel remodelling. Despite the semi-quantitative analysis of the *in-situ* hybridisation presented, further in-depth studies are required to accurately establish a link between *MIR503HG* transcript levels and the extent of remodelling seen in PAH, as well as the contribution of EndMT in the context of human disease. Indeed, recent studies suggested the large contribution of smooth muscle cells and pericytes in the remodelling and lesions observed in PAH[40, 41], ultimately highlighting the multifactorial nature of the disease. It will also be of interest to assess if *MIR503HG* down-regulation is relevant in other vascular remodelling pathologies associated with EndMT, such as atherosclerosis and vein graft failure.

Our data, showing a decrease of mesenchymal markers in mouse ECs after lentiviral delivery of human *MIR503HG* followed by SuHx model of PH, suggests *MIR503HG* plays a key role in the prevention of EndMT. The overexpression of *miR-424/503* in the lung has previously been shown to prevent and rescue the SuHx PH phenotype in rats[42]. This suggests that *MIR503HG* and *miR-424/503* may both contribute to pulmonary hypertension but probably through different processes, supporting a partner relationship between miRNAs and their host gene[43]. Further studies, including normoxic control mice and optimised delivery strategies, are required to address the effect of *MIR503HG* on the severity of the pulmonary hypertension phenotype. At present, the delivery of lncRNA to the EC compartment *in vivo* is too inefficient to justify such approaches.

Taken together, our comprehensive analysis has provided important insights into EndMT and highlighted the critical functional role of the lncRNA *MIR503HG.* Our results open new avenues for targeting EndMT in vascular remodelling, using lncRNA directed therapeutics.

ACKNOWLEGMENTS:
The authors thank G. Aitchison and K. Newton for their technical assistance. Flow cytometry data was generated with support from the QMRI Flow Cytometry and cell sorting facility, University of Edinburgh. Mass Spectrometry data was generated with support from the IGMM Mass Spectrometry facility, University of Edinburgh. We thank the University of Edinburgh Bioresearch & Veterinary Services for exemplary animal husbandry.

SOURCES OF FUNDING:
Andrew H Baker: European Research Council 338991 VASCMIR, BHF Chair and Programme grants (CH/11/2/28733 and RG/14/3/30706) and project grant PG/20/10347. Alex Von Kriegsheim: Wellcome Trust (Multi-user Equipment Grant, 208402/Z/17/Z). Jason Kovacic: US National Institutes of Health (R01HL130423, R01HL135093, R01HL148167-01A1). Matthias S. Leisegang and Ralf P. Brandes: Deutsche Forschungsgemeinschaft (DFG Transregio TRR267, TP A04&A06, EXS2026 Cardiopulmonary Institute - CPI).

DISCLOSURES
None.

SUPPLEMENTAL MATERIALS
Online Data Supplement
Major Resources Table
Online Dataset I EndMT RNAseq
Online Dataset II GoTerms EndMT
Online Dataset III LNT503 RNAseq
Online Dataset IV PD MassSpec
Uncut Gel Blots
References [44-63]

REFERENCES

1. Kovacic JC, Dimmeler S, Harvey RP, Finkel T, Aikawa E, Krenning G, Baker AH. Endothelial to mesenchymal transition in cardiovascular disease: Jacc state-of-the-art review. *J Am Coll Cardiol*. 2019;73:190-209
2. Manavski Y, Lucas T, Glaser SF, et al. Clonal expansion of endothelial cells contributes to ischemia-induced neovascularization. *Circ Res*. 2018;122:670-677
3. Chen PY, Simons M. When endothelial cells go rogue. *EMBO Mol Med*. 2016;8:1-2
4. Dejana E, Hirschi KK, Simons M. The molecular basis of endothelial cell plasticity. *Nat Commun*. 2017;8:14361
5. Zeisberg EM, Tarnavski O, Zeisberg M, et al. Endothelial-to-mesenchymal transition contributes to cardiac fibrosis. *Nat Med*. 2007;13:952-961
6. Hashimoto N, Phan SH, Imaizumi K, Matsuo M, Nakashima H, Kawabe T, Shimokata K, Hasegawa Y. Endothelial-mesenchymal transition in bleomycin-induced pulmonary fibrosis. *Am J Respir Cell Mol Biol*. 2010;43:161-172
7. Evrard SM, Lecce L, Michelis KC, et al. Endothelial to mesenchymal transition is common in atherosclerotic lesions and is associated with plaque instability. *Nat Commun*. 2016;7:11853

8. Cooley BC, Nevado J, Mellad J, et al. Tgf-beta signaling mediates endothelial-to-mesenchymal transition (endmt) during vein graft remodeling. *Sci Transl Med*. 2014;6:227ra234
9. Stenmark KR, Fagan KA, Frid MG. Hypoxia-induced pulmonary vascular remodeling: Cellular and molecular mechanisms. *Circ Res*. 2006;99:675-691
10. Good RB, Gilbane AJ, Trinder SL, Denton CP, Coghlan G, Abraham DJ, Holmes AM. Endothelial to mesenchymal transition contributes to endothelial dysfunction in pulmonary arterial hypertension. *Am J Pathol*. 2015;185:1850-1858
11. Ranchoux B, Antigny F, Rucker-Martin C, et al. Endothelial-to-mesenchymal transition in pulmonary hypertension. *Circulation*. 2015;131:1006-1018
12. Suzuki T, Carrier EJ, Talati MH, Rathinasabapathy A, Chen X, Nishimura R, Tada Y, Tatsumi K, West J. Isolation and characterization of endothelial-to-mesenchymal transition cells in pulmonary arterial hypertension. *Am J Physiol Lung Cell Mol Physiol*. 2018;314:L118-L126
13. Maleszewska M, Moonen JR, Huijkman N, van de Sluis B, Krenning G, Harmsen MC. Il-1beta and tgfbeta2 synergistically induce endothelial to mesenchymal transition in an nfkappab-dependent manner. *Immunobiology*. 2013;218:443-454
14. Kumarswamy R, Volkmann I, Jazbutyte V, Dangwal S, Park DH, Thum T. Transforming growth factor-beta-induced endothelial-to-mesenchymal transition is partly mediated by microrna-21. *Arterioscler Thromb Vasc Biol*. 2012;32:361-369
15. Schmitz SU, Grote P, Herrmann BG. Mechanisms of long noncoding rna function in development and disease. *Cellular and molecular life sciences : CMLS*. 2016;73:2491-2509
16. Monteiro JP, Bennett M, Rodor J, Caudrillier A, Ulitsky I, Baker AH. Endothelial function and dysfunction in the cardiovascular system: The long non-coding road. *Cardiovasc Res*. 2019;115:1692-1704
17. Xiang Y, Zhang Y, Tang Y, Li Q. Malat1 modulates tgf-beta1-induced endothelial-to-mesenchymal transition through downregulation of mir-145. *Cell Physiol Biochem*. 2017;42:357-372
18. Neumann P, Jae N, Knau A, et al. The lncrna gata6-as epigenetically regulates endothelial gene expression via interaction with loxl2. *Nat Commun*. 2018;9:237
19. Muys BR, Lorenzi JC, Zanette DL, et al. Placenta-enriched lincrnas mir503hg and linc00629 decrease migration and invasion potential of jeg-3 cell line. *PLoS One*. 2016;11:e0151560
20. Sorensen I, Adams RH, Gossler A. Dll1-mediated notch activation regulates endothelial identity in mouse fetal arteries. *Blood*. 2009;113:5680-5688
21. Monvoisin A, Alva JA, Hofmann JJ, Zovein AC, Lane TF, Iruela-Arispe ML. Ve-cadherin-creert2 transgenic mouse: A model for inducible recombination in the endothelium. *Dev Dyn*. 2006;235:3413-3422
22. Caruso P, Dunmore BJ, Schlosser K, et al. Identification of microrna-124 as a major regulator of enhanced endothelial cell glycolysis in pulmonary arterial hypertension via ptbp1 (polypyrimidine tract binding protein) and pyruvate kinase m2. *Circulation*. 2017;136:2451-2467
23. Ranchoux B, Harvey LD, Ayon RJ, Babicheva A, Bonnet S, Chan SY, Yuan JX, Perez VJ. Endothelial dysfunction in pulmonary arterial hypertension: An evolving landscape (2017 grover conference series). *Pulm Circ*. 2018;8:2045893217752912
24. Hopper RK, Moonen JR, Diebold I, et al. In pulmonary arterial hypertension, reduced bmpr2 promotes endothelial-to-mesenchymal transition via hmga1 and its target slug. *Circulation*. 2016;133:1783-1794
25. Chen PY, Qin L, Baeyens N, Li G, Afolabi T, Budatha M, Tellides G, Schwartz MA, Simons M. Endothelial-to-mesenchymal transition drives atherosclerosis progression. *J Clin Invest*. 2015;125:4514-4528
26. Goumans MJ, Valdimarsdottir G, Itoh S, Rosendahl A, Sideras P, ten Dijke P. Balancing the activation state of the endothelium via two distinct tgf-beta type i receptors. *EMBO J*. 2002;21:1743-1753

27. Fiedler J, Breckwoldt K, Remmele CW, et al. Development of long noncoding rna-based strategies to modulate tissue vascularization. *J Am Coll Cardiol*. 2015;66:2005-2015
28. Qiu F, Zhang MR, Zhou Z, Pu JX, Zhao XJ. Lncrna mir503hg functioned as a tumor suppressor and inhibited cell proliferation, metastasis and epithelial-mesenchymal transition in bladder cancer. *J Cell Biochem*. 2019;120:10821-10829
29. Fu J, Dong G, Shi H, et al. Lncrna mir503hg inhibits cell migration and invasion via mir-103/olfm4 axis in triple negative breast cancer. *J Cell Mol Med*. 2019;23:4738-4745
30. Chuo D, Liu F, Chen Y, Yin M. Lncrna mir503hg is downregulated in han chinese with colorectal cancer and inhibits cell migration and invasion mediated by tgf-beta2. *Gene*. 2019;713:143960
31. Wang H, Liang L, Dong Q, et al. Long noncoding rna mir503hg, a prognostic indicator, inhibits tumor metastasis by regulating the hnrnpa2b1/nf-kappab pathway in hepatocellular carcinoma. *Theranostics*. 2018;8:2814-2829
32. Huang PS, Chung IH, Lin YH, Lin TK, Chen WJ, Lin KH. The long non-coding rna mir503hg enhances proliferation of human alk-negative anaplastic large-cell lymphoma. *Int J Mol Sci*. 2018;19
33. Lin N, Chang KY, Li Z, et al. An evolutionarily conserved long noncoding rna tuna controls pluripotency and neural lineage commitment. *Mol Cell*. 2014;53:1005-1019
34. Ramos AD, Andersen RE, Liu SJ, et al. The long noncoding rna pnky regulates neuronal differentiation of embryonic and postnatal neural stem cells. *Cell Stem Cell*. 2015;16:439-447
35. Li J, Yang Y, Fan J, Xu H, Fan L, Li H, Zhao RC. Long noncoding rna ancr inhibits the differentiation of mesenchymal stem cells toward definitive endoderm by facilitating the association of ptbp1 with id2. *Cell Death Dis*. 2019;10:492
36. Drasin DJ, Guarnieri AL, Neelakantan D, et al. Twist1-induced mir-424 reversibly drives mesenchymal programming while inhibiting tumor initiation. *Cancer Res*. 2015;75:1908-1921
37. Yan W, Wu Q, Yao W, et al. Mir-503 modulates epithelial-mesenchymal transition in silica-induced pulmonary fibrosis by targeting pi3k p85 and is sponged by lncrna malat1. *Sci Rep*. 2017;7:11313
38. Profumo V, Forte B, Percio S, et al. Leader role of mir-205 host gene as long noncoding rna in prostate basal cell differentiation. *Nat Commun*. 2019;10:307
39. Sun Q, Tripathi V, Yoon JH, et al. Mir100 host gene-encoded lncrnas regulate cell cycle by modulating the interaction between hur and its target mrnas. *Nucleic Acids Res*. 2018;46:10405-10416
40. Steffes LC, Froistad AA, Andruska A, et al. A notch3-marked subpopulation of vascular smooth muscle cells is the cell of origin for occlusive pulmonary vascular lesions. *Circulation*. 2020;142:1545-1561
41. Bordenave J, Tu L, Berrebeh N, et al. Lineage tracing reveals the dynamic contribution of pericytes to the blood vessel remodeling in pulmonary hypertension. *Arterioscler Thromb Vasc Biol*. 2020;40:766-782
42. Kim J, Kang Y, Kojima Y, et al. An endothelial apelin-fgf link mediated by mir-424 and mir-503 is disrupted in pulmonary arterial hypertension. *Nat Med*. 2013;19:74-82
43. Gao X, Qiao Y, Han D, Zhang Y, Ma N. Enemy or partner: Relationship between intronic micrornas and their host genes. *IUBMB Life*. 2012;64:835-840
44. Haeussler M, Schonig K, Eckert H, et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide rna selection tool crispor. *Genome Biol*. 2016;17:148
45. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for crispr screening. *Nat Methods*. 2014;11:783-784
46. Mahmoud AD, Ballantyne MD, Miscianinov V, et al. The human-specific and smooth muscle cell-enriched lncrna smilr promotes proliferation by regulating mitotic cenpf mrna and drives cell-cycle progression which can be targeted to limit vascular remodeling. *Circ Res*. 2019;125:535-551
47. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. Star: Ultrafast universal rna-seq aligner. *Bioinformatics*. 2013;29:15-21

48. Li B, Dewey CN. Rsem: Accurate transcript quantification from rna-seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323
49. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biol*. 2014;15:550
50. Robinson MD, McCarthy DJ, Smyth GK. Edger: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26:139-140
51. Rajkumar AP, Qvist P, Lazarus R, et al. Experimental validation of methods for differential gene expression analysis and sample pooling in rna-seq. *BMC Genomics*. 2015;16:548
52. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36:411-420
53. Pitulescu ME, Schmidt I, Giaimo BD, et al. Dll4 and notch signalling couples sprouting angiogenesis and artery formation. *Nat Cell Biol*. 2017;19:915-927
54. Madisen L, Zwingman TA, Sunkin SM, et al. A robust and high-throughput cre reporting and characterization system for the whole mouse brain. *Nature neuroscience*. 2010;13:133-140
55. Naeije R, D'Alto M. Sex matters in pulmonary arterial hypertension. *Eur Respir J*. 2014;44:553-554
56. Al-Naamani N, Ventetuolo CE. Another piece in the estrogen puzzle of pulmonary hypertension. *Am J Respir Crit Care Med*. 2020;201:274-275
57. Wallace E, Morrell NW, Yang XD, et al. A sex-specific microrna-96/5-hydroxytryptamine 1b axis influences development of pulmonary hypertension. *Am J Respir Crit Care Med*. 2015;191:1432-1442
58. Deng L, Blanco FJ, Stevens H, et al. Microrna-143 activation regulates smooth muscle and endothelial cell crosstalk in pulmonary arterial hypertension. *Circ Res*. 2015;117:870-883
59. Vitali SH, Hansmann G, Rose C, Fernandez-Gonzalez A, Scheid A, Mitsialis SA, Kourembanas S. The sugen 5416/hypoxia mouse model of pulmonary hypertension revisited: Long-term follow-up. *Pulm Circ*. 2014;4:619-629
60. Fehrenbach ML, Cao G, Williams JT, Finklestein JM, Delisser HM. Isolation of murine lung endothelial cells. *Am J Physiol Lung Cell Mol Physiol*. 2009;296:L1096-1103
61. Humbert M, Guignabert C, Bonnet S, et al. Pathology and pathobiology of pulmonary hypertension: State of the art and research perspectives. *Eur Respir J*. 2019;53
62. Baghirova S, Hughes BG, Hendzel MJ, Schulz R. Sequential fractionation and isolation of subcellular proteins from tissue or cultured cells. *MethodsX*. 2015;2:440-445
63. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative pcr and the 2−δδct method. *Methods*. 2001;25:402-408

FIGURE LEGENDS

Figure 1: **EndMT *in vitro* model in venous EC.**
HUVEC were treated with TGF-β2 (10 ng/mL) and/or IL-1β (1 ng/mL) for 7 days. (A) Expression analysis of EndMT markers by RT-qPCR relative to *UBC* (relative quantification:RQ). (*PECAM1* n=5, *ACTA2*, *SNAI2* and *COL1A1* n=8). Statistical analysis was done using a repeated measures one-way ANOVA with Bonferroni correction. (B) Representative images of immunofluorescence staining for PECAM1 (green), SNAI2 (red) and DAPI (blue) (scale bar 50 μm). (C) Quantification of EdU uptake in untreated and treated HUVECs (n=4) with representative FACS histogram plots. (D) Transwell migration assay of treated and untreated HUVECs with representative image of fixed migrated cells stained with DAPI (scale bar 100□M) and quantification (n=3). Statistical analysis of panel C and D was done using linear mixed effects modelling.

Figure 2: single-cell and bulk RNAseq analysis of EndMT identify a common lncRNA signature
(A) tSNE plot of the 5 datasets (untreated HUVEC at day 0, 3 and 7 and TGF-β2 + IL-1β co-treated HUVEC at day3 and day7). **(B)** Violin and tSNE plots of endothelial and mesenchymal signature **(C)** Unsupervised cluster identification **(D)** Violin plot of endothelial (*PECAM1*, *ICAM2*), mesenchymal (*TAGLN*, *COL1A1*), EndMT (*SNAI2*) and proliferation (*MKI67*) marker expression (as z-score) in the different clusters. **(E)** Principal Component Analysis of control and EndMT-induced HUVEC and HPAEC (bulk RNAseq). **(F)** Overlap of EndMT up/down-regulated genes between HUVEC and HPAEC. **(G)** Heatmap of significantly regulated genes and **(H)** of candidate lncRNAs expression (as z-score). **(I)** Violin plot of candidate lncRNA expression in the scRNA-seq clusters.

Figure 3: MIR503HG expression during EndMT *in vitro.*
(A) Schematic representation of the *MIR503HG*, *miRNA-424* and *miRNA-503* loci based on GENCODE v26 annotation. **(B)** Expression of *miRNA-424/-503* in HUVEC ± EndMT (TGF-β2 + IL-1β) (n=4) **(C)** *MIR503HG_2* expression in HSVEC and HCAEC ± EndMT (n=3). **(D)** *MIR503HG_2* expression in HUVEC after treatment with TGF-β2 (50 ng/mL) and H_2O_2 (200nM) (n=3). **(E)** *MIR503HG* correlation to endothelial and mesenchymal signatures based on single-cell RNA-seq. **(F)** *MIR503HG_2*, *18S* and *NEAT1* localisation by cell fractionation. **(G)** Mean *MIR503HG* nuclear copy number compared to unstained negative control (n=4) **(H)** and localisation of *MIR503HG, SNORD3*, and *UBC* by RNA-FISH in HUVEC. Relative quantification of RT-qPCR normalized to *RNU48* or *UBC* relative to Control cells. Statistical analysis was done using linear mixed effects modelling for panel B and C or a repeated measures one-way ANOVA with Bonferroni correction for panel D.

Figure 4: *MIR503HG* knockdown induces EndMT in HUVEC.
(A) *MIR503HG_2* expression in siRNA-mediated *MIR503HG* depletion (si503HG) after 7 days in HUVEC (n=3) **(B)** EndMT marker gene expression in si503HG compared to control after 7 days (n= 3). **(C)** Representative immunofluorescence images of EndMT markers expression in HUVEC (scale bar 50 μm) after knockdown using si503HG. PECAM1 (green), SNAI2 (red) and DAPI (blue). **(D)** *MIR503HG_2* (Left: CRISPR/Cas9-targeted Exon3 (Ex3), right: untargeted Exon1 (Ex1)), **(E)** *miR-424* and *miR-503*, and **(F)** EndMT marker gene expression in HUVEC following CRISPR-mediated deletion of *MIR503HG,* using two lentiviral CRISPR/Cas9 gRNA pairs, after 8 days compared to empty control pairs (n=5). Relative quantification of RT-qPCR normalised to *RNU48* or *UBC* relative to Control cells. Statistical analysis was done using linear mixed effects modelling for panel A and B and a repeated measures one-way ANOVA with Bonferroni correction for panel D-F.

Figure 5: *MIR503HG* overexpression represses EndMT in vitro.
Expression of **(A)** *MIR503HG_2* and **(B)** EndMT marker genes in HUVEC following *MIR503HG* overexpression with LNT_503_2 (MOI 5) with or without TGF-β2 and IL-1β treatment (Control/EndMT)

for 7 days (n=5). Relative quantification of RT-qPCR normalized to *UBC* relative to LNT_CT in Control cells. Analysis by two-way ANOVA with Bonferroni correction. **(C)** Representative immunofluorescence images of EndMT markers in HUVEC following *MIR503HG* overexpression. PECAM1 (green), SNAI2 (red) and DAPI (blue) (scale bar 50 μm). **(D)** Heatmap of the 1683 significant genes between LNT_503HG and LNT_CT in EndMT conditions (displayed as z-score). **(E)** Venn diagram of the overlap between significant changes due to *MIR503HG* overexpression in EndMT-cells (EndMT_LNT_503HG vs EndMT_LNT_CT) and EndMT changes (EndMT vs Control).

Figure 6: *MIR503HG* modulation in mouse is associated with EndMT in PAH
(A) Representative immunofluorescence staining of Normoxia/vehicle or SuHx mouse lung tissue for Cd31 (green), Dapi (blue) and Acta2 or Snai2 (red) (scale bar 50 μm). **(B)** Endothelial and mesenchymal markers expression in TdTomato$^+$ cells isolated from Normoxia/vehicle (n=8) or SuHx mouse lung tissue (n=7). **(C)** *MIR503HG* mouse lncRNA homolog *Gm28730* expression in in TdTomato$^+$ cells isolated from Normoxia/vehicle (n=8 mice) or SuHx mouse lung tissue (n=7). **(D)** Strategy to assess the effect of human *MIR503HG_2* overexpression in EndMT in SuHx PAH model **(E)** Human *MIR503HG_2*, MIR503HG mouse lncRNA homolog *Gm28730* and **(F)** Endothelial/mesenchymal marker expression in CD31$^+$ lung cells isolated from SuHx mouse lung tissue following MIR503HG overexpression with LNT_503HG or LNT_Control (LNT_503HG n=5, LNT_Control n=4). Relative quantification of RT-qPCR normalized to 18S relative to Normoxia (B-C) or LNT_CT (E-F). Statistical analysis of panel B and C was done using an unpaired two-tailed *t*-test except for Col1a1 expression, not following a normal distribution, analysed using a Mann-Whitney test. Panel E and F was analysed using Iman and Conover non-parametric ranking followed by unpaired two-tailed *t*-test.

Figure 7: Human PAH is associated with loss of MIR503HG
(A) *MIR503HG_2* and EndMT markers gene expression in BOECs cells isolated from PAH patients and controls (control n=4, PAH n=5). Relative quantification of RT-qPCR normalized to *UBC* relative to Control cells. Analysis by Iman and Conover non-parametric ranking followed by unpaired two-tailed *t*-test. **(B)** *In situ* hybridization for *MIR503HG* in control and PAH patient lungs, with brightfield staining, and pseudo fluorescence imaging to enhance visualisation of *MIR503HG* (purple/red) and the nucleus (pink/blue). **(C)** Immunofluorescence for smooth muscle cells (ACTA2, green), endothelial cells (vWF, red) and nucleus (DAPI, blue), and IgG controls in control and PAH patient lungs. Scale bar 50□m. Dotted squares denote high-power view of vessels.

Figure 8: *MIR503HG* binding partner PTBP1 regulates EndMT
(A) Schematic of *MIR503HG in-vitro* pulldown experiment. **(B)** Dot plot of -log10 pvalue versus log2 fold change of *MIR503HG_2* interacting proteins compared to GFP control after RNA pulldown assay and mass spectrometry (n=3) **(C)** RT-qPCR analysis of *MIR503HG_2* and *UBC* in RNA immunoprecipitation of IgG, PTBP1 and HNRNPA0 from HUVEC lysate (n=3). **(D)** Western blot analysis of PTBP1 level in si503HG in HUVEC after 7 days. GAPDH used as a loading control. **(E)** Western blot analysis of PTBP1 level in CONTROL and EndMT HUVEC. GAPDH used as a loading control **(F)** EndMT marker expression in siPTBP1 samples 7 days after transfection (n=5). **(G)** Overlap of genes significantly down-regulated by LNT_503 in EndMT condition, up-regulated in EndMT compared to control and up-regulated by PTBP1 in HepG2 cells. Relative quantification of RT-qPCR normalized to *UBC* relative to siCONTROL treatment. Statistical analysis was done using a repeated measures one-way ANOVA with Bonferroni correction for panel C and F and linear mixed effects modelling for panel D-E.

NOVELTY AND SIGNIFICANCE

What Is Known?

- Endothelial to mesenchymal transition (EndMT) is a complex process contributing to vessel remodeling during disease, including Pulmonary Arterial Hypertension (PAH).
- Long non-coding RNAs (lncRNAs) regulate many biological processes and are involved in vascular biology and disease.

What New Information Does This Article Contribute?

- Transcriptomic changes associated with EndMT include lncRNAs.
- The lncRNA *MIR503HG* is downregulated during EndMT *in vitro*, as well as in mouse and human PAH samples.
- Down-regulation of *MIR503HG* level induces a spontaneous EndMT profile which may be mediated, in part, by *MIR503HG* interaction with the RNA binding protein PTBP1.

EndMT transition occurs during development and pathological vessel remodeling. Regulators of EndMT are still poorly characterized. LncRNAs are key regulators of many biological processes and involved in disease. Here, we identify lncRNAs regulated during EndMT and show that the lncRNA *MIR503HG* expression is decreased during EndMT *in vitro*. Loss of *MIR503HG* was also observed both in mouse and human PAH samples. Down-regulation of *MIR503HG* alone recapitulates EndMT while its overexpression reduces EndMT. MIR503HG mechanism of action was uncovered, by identifying the RNA binding protein PTBP1 as a binding partner and showing PTBP1 regulation of mesenchymal genes. We provide early evidence of *MIR503HG* therapeutic potential in a mouse model.

FIGURE 1